Ultra Wide Band Modulation

Ultra Wide Band (UWB) modulation techniques are the latest buzzword to hit the industry press. In this feature we will explore some of the ideas behind these interesting techniques, and attempt to place them into context.

Most readers will be familiar with established narrowband and spread spectrum modulation schemes. To best contrast how different UWB techniques are, it is worth briefly reiterating the basic principles of narrowband and spread spectrum techniques.

Narrowband modulation techniques are the oldest and technologically simplest approach, and have also set the precedents via which bandwidth is allocated commercially. Whether you are or are not a fan of narrowband techniques, they remain the backbone of modern wireless communications and form the basis of every regulatory model in existence. The idea behind narrowband modulation is in principle very simple. You take a sine shaped carrier wave and you change its amplitude, phase, or both to impose a signal.

The simplest technique is amplitude modulation, the basis of AM radio and analogue TV. The instantaneous amplitude of the carrier wave is proportional to the instantaneous magnitude of the modulation signal. Modulation is trivial to produce, and trivial to extract, but also trivial to corrupt with an interfering signal, spike or noise.

The alternative is to use either phase or frequency modulation techniques, whereby the instantaneous phase or frequency of the carrier is proportional to modulation signal. FM or PM schemes have the very nice property of being more difficult to interfere with, since additive noise and signals are more easily rejected by the demodulation hardware. FM is the basis of FM radio and analogue TV sound channels (and colour in the SECAM standard).

The final, and least used modulation scheme, is time based modulation, whereby the carrier wave is pulsed, and the pulses shifted in time to reflect the modulation signal amplitude. Pulse Position Modulation (PPM) was experimented with but never found serious commercial uses.

More sophisticated narrowband schemes have become widely used in recent years, based largely upon variations on the theme of Quadrature Amplitude Modulation (QAM), whereby the modulation signal is imposed by varying both phase and amplitude. This is the technique pioneered in NTSC and PAL-A/B colour subcarriers, and is the basis of most modems, many cable modems, and with further variations, the US HDTV transmission standard. Whether the modulation is analogue, or a streams of ones and zeros, many possible narrowband schemes are available.

The one attribute they all share in common, to greater or lesser degrees, is a very small ratio of modulation envelope bandwidth against the frequency of the carrier itself.

Shannon's information theory shows us that the ability of a modulation scheme to resist noise and interference depends strongly upon how much bandwidth is used to carry the information. The wider the modulation bandwidth in relation to the modulation signal bandwidth, the more resilient the signal becomes.

This idea did not go unnoticed, and by the 1950s spread spectrum modulation schemes began to appear, intended to exploit this important idea.

The simplest spread spectrum technique is Direct Spreading (DS), whereby the carrier wave is amplitude or phase modulated with a pseudorandom (pseudo-noise or PN) sequence. A typical arrangement is where a one or a zero is represented by a PN sequence, or its inverse. The larger the number of bits (chips) in the PN sequence, the greater the modulation bandwidth or spreading ratio and the more resilient the modulation.

The basic alternative to DS is frequency hopping (FH), whereby the carrier wave is pseudo randomly hopped in frequency, in a manner analogous to direct spreading. As with direct spreading, the larger the number of slots across which the carrier is hopped, the greater the spreading ratio and resilience of the modulation.

Spread spectrum techniques have the very nice property of being able to share bandwidth between multiple signals concurrently, by using different PN codes for modulation. Providing these codes have the important mathematical property of mutual orthogonality in a cross-correlation operation, then multiple signals can occupy the same bandwidth. This is called Code Division Multiple Access or CDMA. There are of course no free lunches here, in the sense that only so many orthogonal codes may be used concurrently before mutual interference arises.

Security is also improved, since without knowing the spreading code, you cannot demodulate the signal. The longer the spreading code, the harder it is to guess it. Another nice property of spread spectrum techniques is that the power density per bandwidth is much lower than in narrowband schemes, in proportion to the spreading ratio. With a very large spreading ratio the signal spectrum becomes distinctly noise-like.

Both DS and FH are incorporated in the 802.11 wireless LAN standards package, while DS is the basis of CDMA mobile telephony and GPS satellite navigation.

As with narrowband schemes, hybrids have also become very popular. The military JTIDS datalink, the backbone of US and NATO battlefield digital networking, uses a combination of DS and FH techniques to produce a highly jam resistant signal.

Much of the advantage of spread spectrum techniques over narrowband techniques stems from their much greater ratio of modulation bandwidth to signal bandwidth. If we consider also the advantages of CDMA techniques, there are important gains to be made from using spread spectrum over narrowband.

Despite this spread spectrum techniques have been slow to penetrate into the vast commercial marketplace, with the most important inroads only made in wireless LANs. The military have made greater gains, but only through the need to deal with eavesdropping and jamming.

The great technological enabler for spread spectrum techniques has been the Monolithic Microwave IC (MMIC), which enables the complex and fast circuitry needed for spreading modulators and receiver correlators to be manufactured affordably in large volumes. Indeed, until recently the cost barriers were the reason why spread spectrum was used mostly in military designs.

Narrowband techniques still rule the world, but the advantages of spread spectrum will see it proliferate in coming years to occupy increasing portions of the total market. Whether it manages to ever wholly displace narrowband techniques remains to be seen. The current regulatory models for allocating bandwidth have been structured around narrowband techniques and hungrily carved up by broadcasters and other users. Spread spectrum becomes prohibitively expensive to buy bandwidth for in such a regime, despite its obvious advantages, and requires a different regime for bandwidth allocation, one in which codes in a shared band are sold, rather than chunks of the band itself.

Enter UWB techniques, which further extend the ideas used in spread spectrum.

Ultra Wide Band Modulation

UWB techniques aim to increase the spreading ratios used quite dramatically over established spread spectrum techniques, thereby exploiting Shannon's model to a greater degree.

Narrowband techniques are inherently a carrier centred approach, indeed in most such modulations the carrier wave itself is discrete and is used as a reference signal for demodulating the modulation signal. Spread spectrum techniques are also tied to a carrier wave, even if almost all of the energy of the signal lies in the modulation sidebands carrying the signal.

UWB techniques radically depart from this orthodoxy, as they do not employ a carrier wave in the conventional sense.

Probably the best conceptual starting point for UWB is to explore that favourite plaything of electrical engineers, the Dirac impulse. Infinitely short, the impulse has an infinite bandwidth. It is a mathematical abstraction which we cannot replicate in the physical world.

However, if we produce an extremely short pulse, in the frequency spectrum it will spread its energy over a very large bandwidth, which gets wider as the pulse gets shorter.

If we are aiming to exploit Shannon, we want as wide a bandwidth as possible. The question which arises is of course, how to produce a practical modulation scheme which can exploit the properties of very short pulses, yet be implementable with affordable hardware.

Several obvious problems arise immediately:

- What kind of signal do we use to emulate the behaviour of an impulse ?

- How do we encode the modulation on to the UWB signal ?
- How to make a wideband antenna which is directional ?
- How to make the modulation hardware to produce UWB signals ?

The first question is interesting indeed. Clearly, the modulation scheme has to put out a very short burst of energy, followed by a big gap, to acquire the huge spreading ratio required.

Current researchers are exploring four techniques for doing this:

- Singlet impulses, essentially a spike with a DC component.
- Doublets, or paired impulses.
- Monocycles, essentially a single cycle of a sinewave which has been shaped.
- Wavelets, essentially a small number of cycles of a sinewave, which have been shaped.

Wavelets and monocycles appear to be the favourites in the UWB literature, since techniques exist for generating them and the mathematics behind them are reasonably tractable. Being able to produce a monocycle or a wavelet does not however answer the question of how to encode the message.
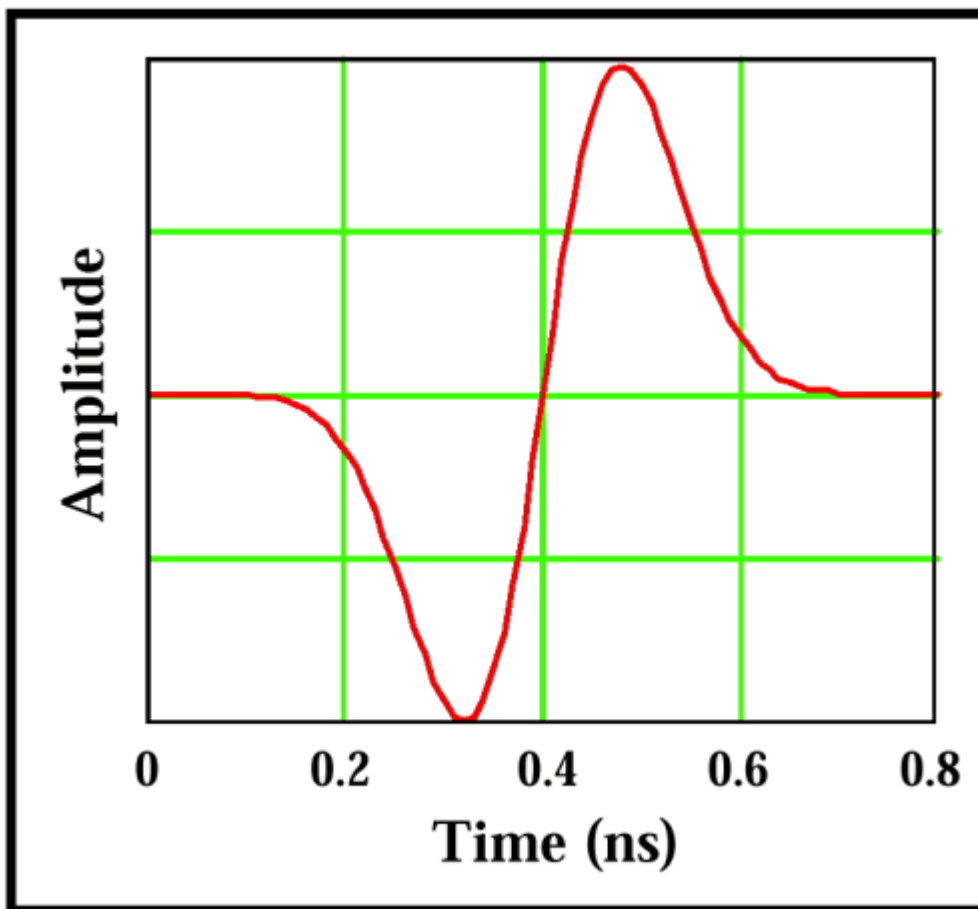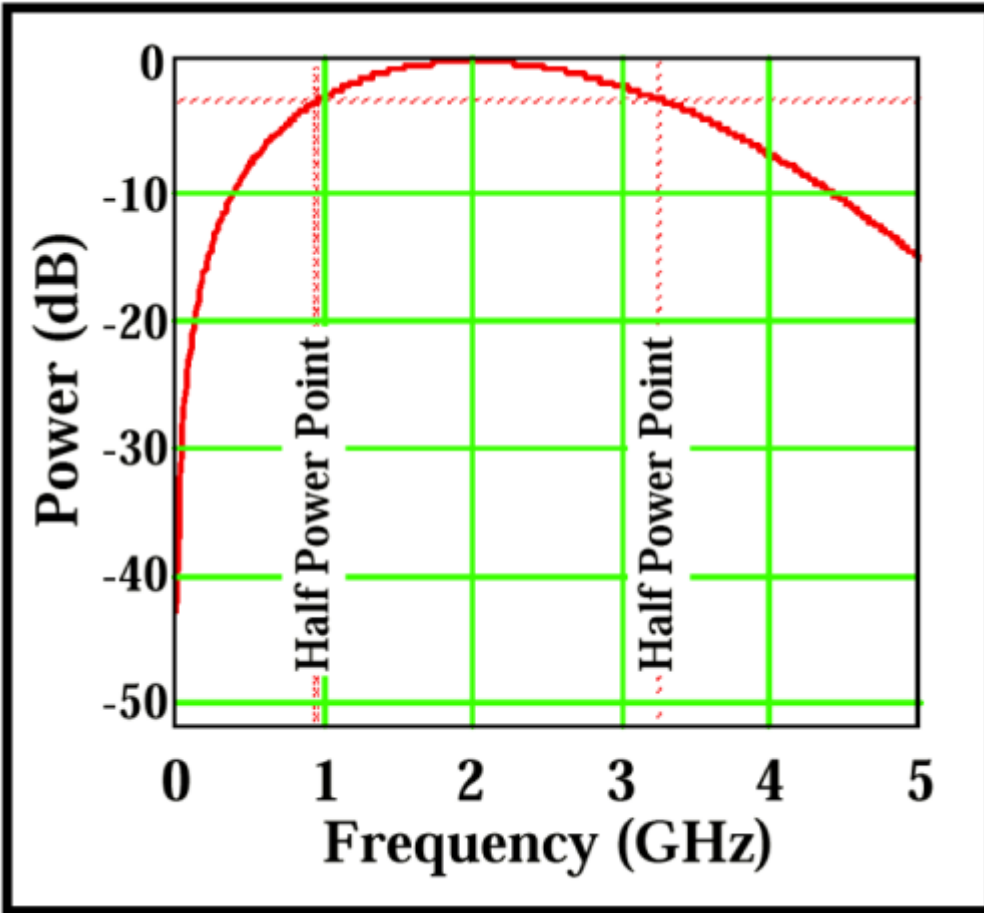
A monocycle and its spectrum is depicted in Figures 1 and 2.

The encoding scheme which has received the most coverage to date is the Time Modulated UWB technique, pioneered by Time Domain (http://www.time-domain.com) in the US.

This technique sees a stream of monocycles produced at a constant average rate, each of which is extremely short in relation to the gaps between them.

Modulation is then imposed by shifting the monocycles in time, to arrive either before or after a datum point in time, advanced or delayed by a fixed increment in time. This is a direct equivalent to the well known but infrequently used Pulse Position Modulation (PPM) scheme, indeed it differs from the conventional primarily in the fact that a burst of carrier wave has been replaced by a monocycle.

In a simple arrangement, an early pulse signals a 1 and a late pulse a 0. This indeed can be used for transmitting a digital message. What happens if two links try to occupy the same bandwidth ? Obviously they interfere.

This is where some clever thinking has been applied by Time-Domain, and their supporting researchers. Borrowing an important idea from spread spectrum technology, they use a pseudorandom time delay rather than a fixed one. In this manner, they combine the properties of a spread spectrum signal, but achieve a much higher spreading ratio by using monocycles rather than a continuous carrier wave.

Termed time hopping impulse modulation or time hopping spread spectrum modulation, this technique combines the attributes of conventional spread spectrum with those of impulse PPM. Figures 3 and 4 depict the modulation scheme.

By using a PN code modulation, and a correlation receiver, Time Hopping Spread Spectrum (THSS) acquires the same capability for code division multiplexing seen in established spread spectrum techniques. The big difference lies in the additional spreading effect produced by using a pulse modulation rather than carrier wave modulation.

Implementation of hardware to perform THSS is not a trivial chore by any measure.

The PulsON product currently being promoted by Time Domain Corp uses a 16 bit delay word, to control the timing of a 500 picosecond (0.5 nanosec) monocycle, in a 100 nanosecond frame, with an accuracy of 10 picoseconds, and required linearity in timing of 10 picoseconds. Figure 5 shows the system block diagram.

The chips containing the modulation hardware and correlation hardware for the receiver had to be implemented in SiGe heterojunction MMIC technology, to be capable of achieving such an ambitious specification. A third chip is in development to provide the low speed functions required for a complete low cost receiver.

The signal modulation bandwidths published are of the order of 2 GHz or higher, resulting in spreading ratios of 50 dB for a 20 kilobit/s data rate.

To place this into context, the spreading ratio of 802.11 WLAN is around 10 dB, differing by a factor of 10,000 !

The enormous spreading ratio has two important implications:

- The process gain is so great that interfering narrowband signals can have power levels commensurately higher. The practical consequence is very lower power output for the THSS UWB system.
- The required antenna bandwidth is huge by conventional standards, making most directional antenna types unusable for UWB. The practical consequence has been a focus upon omni aerials for use in local area applications.

Other interesting consequences follow from THSS UWB.
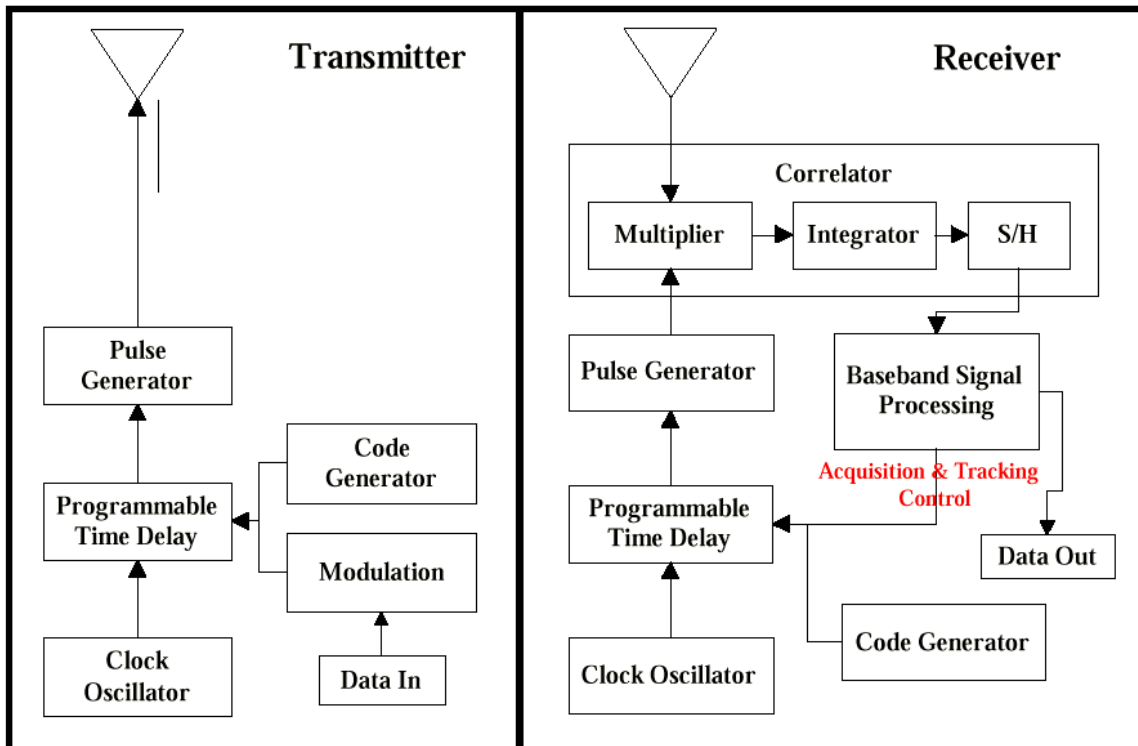
Perhaps the most important one is the issue of multipath propagation, a genuine plague in mobile telephony and WLANs, especially indoors. Multipath arises when receivers see multiple, time delayed copies of the same transmission, produced by reflections from objects along the transmission path. The carrier wave seen by the receiver is a sum of these components, which interfere and result in fading. Often such fading can render a link unusable.

A UWB system using monocycles or wavelets does not suffer this problem. This is for the simple reason that the monocycle or wavelet occupies a length in free space of the order of inches. Consecutive monocycles or wavelets are thus separated in space by tens to a hundred feet, and any reflected examples will mostly arrive well after the directly received monocycle or wavelet. Since a PN code is being used, the correlator will see them as noise and ignore them.

The strategy adopted by Time Domain is common to established spread spectrum designs, using a rake receiver with multiple correlators. This ensures that should a destructive cancellation arise due to multipath, the receiver can find another signal to demodulate.

Radar and Range finding Applications

The monocycle THSS scheme has numerous applications other than telephony and WLANs. The very accurate timebase and short duration of the pulses mean that short range GPS-like rangefinding and positioning schemes with accuracies under a centimetre are feasible. Indeed, Time Domain Corp papers elaborate on this in some detail.

Such technology has a huge range of industrial applications, as well as automotive applications. Consider the joys of tight parking!

Radar applications have also been studied intensively, both for commercial, law enforcement, burglar alarm and military applications. Some very ambitious proposals have been floated for UWB radar, including the ability to image the shape of an object, the ability to detect stealth aircraft and the ability to produce ultra high resolution synthetic aperture radar imagery.

While the short range radar applications are feasible in the near term, antenna bandwidth issues would suggest that any of the more ambitious ideas will have to wait for some time yet, until antenna technology can catch up.